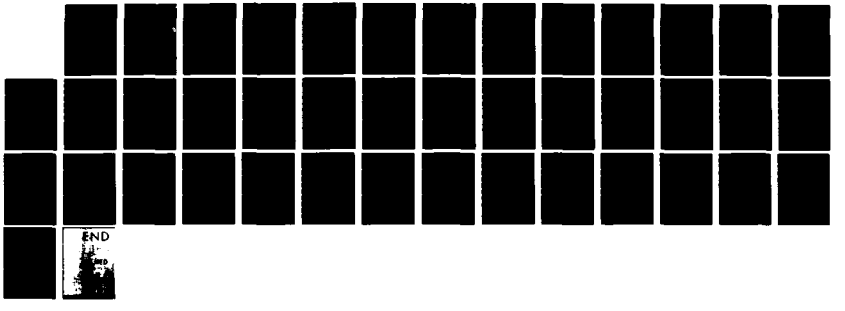
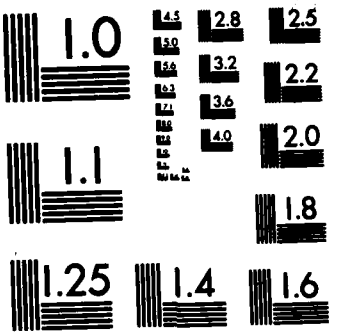
AD-A126 899 A BAYESIAN APPROACH TO THE DESIGN OF DECISION RULES FOR FAILURE DETECTION..(U) MASSACHUSETTS INST OF TECH CAMBRIDGE LAB FOR INFORMATION AND D.. E Y CHOW ET AL. 1/1

UNCLASSIFIED 14 FEB 83 LIDS-P-1281 N00014-77-C-0224 F/G 12/1 NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

February 14, 1983
To be submitted to IEEE Transactions on Aerospace and Electronic Systems

LIDS-P-1281

AD A126899

A Bayesian Approach to the Design of Decision Rules for Failure Detection and Identification*

Edward Y. Chow

Schlumberger-Doll Research
P. O. Box 307
Ridgefield, CT 06877
Tel. (203) 431-5419

and

Alan S. Willsky

Laboratory for Information and Decision Systems
Massachusetts Institute of Technology
Cambridge, MA 02139
Tel. (617) 253-2356

DTIC ELECTE APR 12 1983 D H

Abstract

The formulation of the decision making process of a failure detection algorithm as a Bayes sequential decision problem provides a simple conceptualization of the decision rule design problem. As the optimal Bayes rule is not computable, a methodology that is based on the Bayesian approach and aimed at a reduced computational requirement is developed for designing suboptimal rules. A numerical algorithm is constructed to facilitate the design and performance evaluation of these suboptimal rules. The result of applying this design methodology to an example shows that this approach is potentially a useful one.

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

DTIC FILE COPY

* This work was supported in part by the Office of Naval Research under Contract No. N00014-77-C-0224 and in part by NASA Ames Research Center under Grant No. NGL-22-009-124.

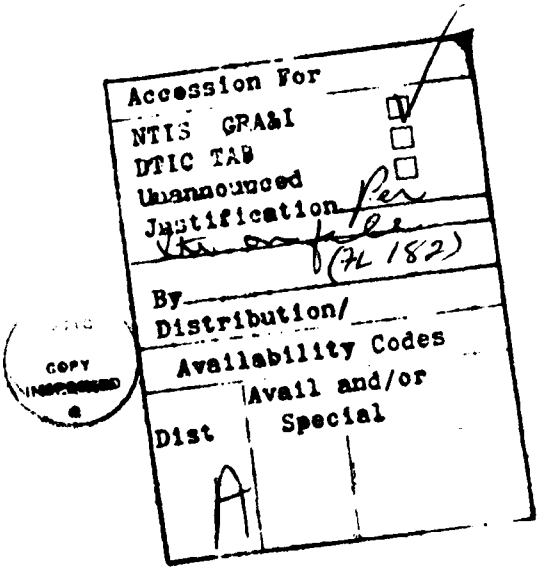
83 04 08 095

1. INTRODUCTION

A failure detection and identification (FDI) process consists of two basic stages : residual generation and decision making. In the first stage, sensor outputs are processed to form residuals that typically have distinct characteristics under normal (no-fail) conditions and under the various possible failures modes. (See [1] for a discussion of the design of residual generation processes.) The function of the second stage is to monitor the residuals and make decisions concerning the occurrence and identity of failure modes. The decision mechanism is based on a compromise among speed of detection, false alarm rates, and identification accuracy, and it belongs to the extensively studied class of sequential tests or sequential decision rules [2-15]. Most previous works, however, were focussed on either the detection of a single type of change (failure) [5-9], or the sequential testing of M hypothesis, which is analogous to the problem of identifying the failure mode given the onset time is known [12-14]. In this paper, we employ the Bayesian approach to the design of decision rules that *directly confront the problem of detecting and distinguishing the various possible failure modes which may occur at unknown times.*

In Section 2 we describe the Bayes formulation of the FDI decision problem. Although the optimal rule is generally not computable the structure of the Bayesian approach can be used to derive practical suboptimal rules. The design of suboptimal rules based on the Bayes formulation is discussed in Section 3. The approximations and simplifications that are made in order to obtain these rules make systematic use of the important features specific to the problem of dynamic failure detection and consequently allow us to interpret each step in our simplification procedure in terms of

its implications for failure detection. In Section 4 we report on our experience with this approach to designing decision rules through a numerical example and simulation.



2. THE BAYESIAN APPROACH

In this section we adapt and specialize the standard Bayes Sequential Decision Problem (BSDP) [15] to the problem of failure detection. The BSDP formulation of the FDI problem consists of six elements:

1. Θ : the set of states of nature or failure hypotheses. For simplicity in this development we assume that only single failures may occur. In general an element θ of Θ conveys several pieces of information, namely, the type of failure mode, its time of occurrence, and probably a variable specifying the severity of the failure. For example, if a particular failure mode corresponds to the onset of a sensor bias, the level of this bias could be specified in the corresponding element of θ. In many applications, however, it suffices simply to identify the failure type without estimating its severity. Furthermore, what is often done to eliminate this nuisance parameter completely is to hypothesize a fixed scale for each failure type corresponding to the smallest deviation from normal behavior that one would like to detect. For example, this approach was used with great success for the detection of aircraft sensor failure in [16]. We will adapt this approach here, and consequently elements of θ are 2-tuples, $\theta = (i,\tau)$, corresponding to the onset of the ith failure mode at time τ. We assume that there are M hypothsized failure modes and also denote by (0,-) that element of Θ corresponding to no failure. Thus,

$$\Theta = \{(i,\tau),\ i = 1,...M,\ \tau = 1,2,...\} \cup \{(0,-)\}$$

2. μ : the prior probability mass function (PMF) over the nature set Θ. This PMF

represents the a priori information concerning possible failures, i.e. how likely it is for each type of failure to occur, and when is a failure likely to occur. Because this information may not be available or accurate in some cases, the need to specify μ is a drawback of the Bayes approach for such cases. Nevertheless, we will see that it can be regarded as a design parameter in the specification of the Bayes rule.

In general, μ may be arbitrary. Here, we assume the underlying failure process has two properties: 1) the occurrence of each of the M failure modes is independent of the other, and 2) the occurrence of each failure i is a Bernoulli process with (success) parameter ρ_i, a common model for failure s in physical components. The independent assumption is also a reasonable one in most applications. It is straightforward to show that

$$\mu(i,\tau) = a(i)\rho(1-\rho)^{\tau-1} \qquad i = 1,...,M,\ \tau = 1,2,...$$

where

$$\rho = 1 - \prod_{j=1}^{M}(1-\rho_j)$$

$$a(i) = \rho_i(1-\rho_i)^{-1}\left[\sum_{j=1}^{M}\rho_j(1-\rho_j)^{-1}\right]^{-1}$$

The parameter ρ may be regarded as the parameter of the combined (Bernoulli) failure process which specifies the statistics of the occurrence of the first failure; a(i) can be interpreted as the marginal probability that the first failure is of type i. Note that the present choice of μ indicates that the arrival of the first failure is memoryless.

3. $D(k)$: The discrete set of terminal decisions available to the decision maker when the residual-monitoring is interrupted at time k in order to make failure identification. An element δ of $D(k)$ may denote the pair (j,t), i.e. the declaration of a type i failure to have occurred at time $t \leq k$. Alternatively, δ may represent an identification of the j-th failure without regard for the failure time, or it may signify the presence of a failure without specifying its type or time, i.e. simply an alarm. Note that the number of terminal decisions specifying failure times grow with k (as there are more times at which a failure could have occurred) while the number of decisions not specifying times will remain the same. In addition, $D(k)$ does not include the declaration of no-failure, since the residual monitoring is stopped only when a failure appears to have occurred. It is worth pointing out that in some application one may not be interested in estimating failure onset times, there are others in which one is. For example, if a failed sensor has been used for sometime in a closed-loop filter and control law, one may wish to estimate how long the failure has been present in order to compensate for the effect of this erroneous signal. In addition, onset time estimates are critical in other event detection problems such as electrocardiogram analysis [17] and maneuver detection [18,19].

4. $L(k;\theta,\delta)$: the terminal decision cost function at time k. $L(k;\theta,\delta)$ denotes the penalty for deciding $\delta \in D(k)$ at time k when the true state of nature is θ. It is assumed to be bounded and non-negative and have the structure :

$$L(k;\theta,\delta) = \begin{cases} L((i,\tau),\delta), & \tau \leq k,\ \delta \in D(k) \\ L_F, & \tau > k,\ \delta \in D(k) \end{cases}$$

$$L(k;(0,-),\delta) = L_F$$

where $L((i,\tau),\delta)$ is the underlying cost function for deciding δ when failure (i,τ) has already occurred. Also, L_F denotes the penalty for a false alarm (note that a false alarm corresponds to making a failure declaration before one occurs), and it can be generalized by allowing it to be a function of δ.

The cost function $L((i,\tau),\delta)$ generally has some additional structure. For example, a terminal decision cost that indicates the correct failure (and/or onset time) should receive a lower cost then one with the wrong failure (and/or onset time) indication. We further assume that the penalty due to an incorrect identification of the failure time is only dependent on the error in such an identification. That is for $\delta = (j,t)$,

$$L(k;(i,\tau),(j,t)) = L(i,j,(t-\tau))$$

Note that $L(i,i,(t-\tau))$ corresponds to the penalty for an incorrect time estimate of when the failure type is correctly determined. Again the use and importance of this cost depends upon the application. Finally, if onset time is unimportant, so that δ does not obtain a time specification, we have

$$L((i,\tau),\delta) = L(i,\delta)$$

5. $r(k)$: the m-dimensional residual (observation) sequence. We shall let $p(r(1),\ldots r(k)\,|\,i,\tau)$ denote their joint conditional density when (i,τ) is true. Since the residual is affected by the failure in a causal manner, its conditional density has the property

$$p(r(1),...,r(k)|i,\tau) = p(r(1),...,r(k)|0,-), \quad i=1,...,M,\ \tau>k$$

In this paper, we will assume that the residual is an independent Gaussian sequence with V (mxm matrix) as the time-independent covariance function and $g_i(k-\tau)$ as the mean given that the failure (i,τ) has occurred. With the covariance assumed to be the same for all failures, the mean function $g_i(k-\tau)$ characterizes the effect of the failure (i,τ), and it is henceforth called the signatures of (i,τ) (with $g_i(k-\tau)=0$ for $i=0$, or $\tau \geq k$). We have chosen to study this type of residuals because its special structure facilitates the development of insights into the the design of decision rules. Such a model arises in the case in which the residuals are generated by a Kalman filter based on normal operation and in which the failure enter additively in the system dynamics or sensor outputs [20]. While this model is not correct if parametric failure are considered (since in this case the correlation structure of the residuals is also affected by the failure), the general concepts we develop for the formulation of a BSDP for failure detection carry over to the parametric case. Furthermore, as reported in [16,21], an FDI system based on an appropriate additive-failure model can often work very well in detecting parametric failures.

6. $c(k,(i,\tau))$: the delay cost function having the properties :

$$c(k,(i,\tau)) = \begin{cases} c(i,k-\tau)>0, & \tau<k \\ 0, & \tau \geq k \end{cases}$$

$$c(i,k_1-\tau) > c(i,k_2-\tau), \qquad k_1>k_2>\tau$$

After a failure has occurred at time τ, there is a penalty for delaying the terminal

decision until time $t \geq \tau$ with the penalty an increasing function of the delay $(k-\tau)$. In the absence of a failure, no penalty is imposed on residual sampling. In this paper we will consider a delay cost function that is linear in the delay, i.e. $c(i,k-\tau) = c(i)(k-\tau)$, where $c(i)$ is a positive function of the failure type i, and may be used to provide different delay penalty for different types of failures.

A sequential decision rule naturally consists of two parts : a stopping rule (sampling plan) and a terminal decision rule. The stopping rule is essentially a detection rule as its purpose is to determine whether monitoring should be interrupted in order to identify a failure. The terminal decision rule then performs the subsequent identification. The stopping rule denoted by $\Phi = (\phi(0),\phi(1;r(1)),...,\phi(k;r(1),...,r(k)),...)$ is a sequence of functions of the observed residual samples, with $\phi(k;r(1),...,r(k)) = 1$ or 0. When $\phi(k;r(1),...,r(k)) = 1$ (0), residual-monitoring or sampling is interrupted (continued) after the k-th residual sample, $r(k)$, is observed. Alternatively, the stopping rule may be defined by another sequence of functions $\Psi = (\psi(0),\psi(1;r(1)),...,\psi(k;r(1),...,r(k)),...)$, where $\psi(k,r(1),...r(k)) = 1$ indicates that residual-monitoring has not been interrupted up to and including time (k-1) but will be interrupted when residual samples $r(1),...,r(k)$ are observed [15]. The functions Φ and Ψ are related to each other in the following way :

$$\psi(k;r(1),...r(k)) = \phi(k;r(1),...,r(k)) \prod_{s=0}^{k-1} [1-\phi(r(1),...,r(s))], \quad k \geq 1$$

with $\psi(0) = \phi(0)$.

The terminal decision rule is a sequence of functions, $D=(d(0),d(1;r(1)),...,d(k;r(1),...,r(k)),...)$. The function $d(k;r(1),...r(k))$ maps the

residual samples r(1),...,r(k) into the terminal decision set $D(k)$ and represents the decision rule used to arrive at a failure identification if sampling is interrupted at time k.

If (i,τ) is the true state of nature and if the sequential decision rule (Φ,D) is used, then the total expected cost, i.e. the expectation of the sum of the delay and terminal decision costs is

$$U[(i,\tau),(\Phi,D)] = \sum_{k=0}^{\infty} E_{i,\tau}\{\psi(k;r(1),...,r(k))\,[c(k,(i,\tau)) + L(k;(i,\tau),d(k;r(1),...,r(k)))]\}$$

where $E_{i,\tau}$ denotes the expectation given that (i,τ) is true. The Bayes Sequential Decision Rule (BSDR) with respect to μ is defined to be the sequential decision rule (Φ^*,D^*) that minimizes the sequential Bayes risk $U_s(\Phi,D)$ which is given by

$$U_s(\Phi,D) = E\{U[(i,\tau),(\Phi,D)]\}$$

$$= \sum_{i=1}^{M}\sum_{\tau=1}^{\infty} \mu(i,\tau)U[(i,\tau),(\Phi,D)]$$

Now we discuss an interpretation of the sequential Bayes risk for the FDI problem. Let us define the following notation

$$P_F(\tau) = \sum_{k=1}^{\tau-1} E_{0,-}\{\psi(k;r(1),...,r(k))\}$$

$$D = \bigcup_{k=0}^{\infty} D(k)$$

$$\tilde{S}(k,\delta) = \{[r(1),...,r(k)]: \psi(k;r(1),...,r(k))=1,\ d(k;r(1),...,r(k))=\delta\},\quad \delta \in D$$

$$\Pr\{\tilde{S}(k,\delta)\,|\,i,\tau\} = \int_{\tilde{S}(k,\delta)} p(r(1),\ldots,r(k)\,|\,i,\tau)\,dr(1),\ldots,dr(k)$$

$$\bar{t}(i,\tau) = \sum_{k=\tau}^{\infty} (k-\tau)\,(1-P_F(\tau))^{-1} E_{i,\tau}\,\{\psi(k;r(1),\ldots,r(k))\}$$

$$P((i,\tau),\delta) = (1-P_F(\tau))^{-1} \sum_{k=\tau}^{\infty} \Pr\{\tilde{S}(k,\delta)\,|\,i,\tau\}$$

where $P_F(\tau)$ is the probability of stopping to declare a failure before the failure occurs at τ, i.e. the probability of false alarm when a failure occurs at time τ or later. D is the set of terminal decisions for all times. $\tilde{S}(k,\delta)$ is the region in the sample space of the first k residuals where the sequential rule (Φ,D) yields the terminal decision δ. Clearly, the $\tilde{S}(k,\delta)$'s are disjoint sets with respect to both k and δ. The expressions $\bar{t}(i,\tau)$ and $P((i,\tau),\delta)$ are respectively the conditional expected delay and the conditional probability of declaring δ, given a type i failure has occurred at time τ and no false alarm has been signalled before this time. $P((i,\tau),\delta)$ is called the generalized cross-detection probability. Using these quantities the sequential Bayes risk can be written as

$$U_s(\Phi,D) = \sum_{i=1}^{M} \sum_{\tau=1}^{\infty} \mu(i,\tau)\,\{L_F P_F(\tau) + (1-P_F(\tau))\,[c(i)\bar{t}(i,\tau) + \sum_{\delta \in D} L((i,\tau),\delta) P((i,\tau),\delta)]\} \qquad (1)$$

Equation (1) indicates that the sequential Bayes risk is a weighted combination of the conditional false alarm probability, expected delay to decision and cross-detection probabilities, and the optimal sequential rule (Φ^*,D^*) minimizes such a combination. From this vantage point, the cost functions (L and c) and the prior distribution (μ) act as the weighting coefficients and hence serve as a basis for specifying the tradeoff relationships among the various performance issues. The advantage of this approach

is that only the total expected cost instead of every individual performance issue needs to be considered explicitly in designing a sequential rule. The drawback, however, lies in the need to choose a set of appropriate cost functions (and the prior distribution) when the physical problem does not have a natural set, as it doesn't in general. In this case, the Bayes approach is most useful with the cost functions and the prior distribution considered as design parameters that may be adjusted to obtain an acceptable design.

The optimal terminal decision rule D^* can be easily shown to be a sequence of fixed-sample-size tests [15]. The determination of the optimal stopping rule Φ^* is a dynamic programming problem [22]. The immense storage and computation required make Φ^* impossible to compute, and suboptimal rules must be used.

Despite the impractical nature of its solution, the BSDP provides a useful framework for designing suboptimal decision rules for the FDI problem because of its inherent characteristic of explicitly weighing the tradeoffs between detection speed and accuracy (in terms of its cost structure). A sequential decision rule specifies a set of sequential decision regions $\tilde{S}(k,\delta)$, and the decision regions corresponding to the BSDR yields the minimum risk. From this vantage point, the design of a suboptimal rule can be viewed as the problem of choosing a set of decision regions that would yield a reasonably small risk. This is the essence of the approach to suboptimal rule design that we take in this paper and describe next.

3. DESIGN OF SUBOPTIMAL RULES

3.1 Suboptimal Rules Based on the BSDR

The Sliding Window Approximation

The immense computation associated with the BSDR is partly due to the increasing number of possible failure times that must be considered as time progresses. The remedy for this problem is the use of a sliding window to limit the number of failure onset times to be considered at each time. The assumption made under the sliding window approximation is that essentially all failures can be detected within W time steps after they have occurred, or that if a failure is not detected within this time it will not be detected in the future. Here, the window size W is a design parameter, and it should be be chosen long enough so that detection and identification of failures are possible, but short enough so that implementation is feasible [22].

The sliding window rule (Φ^W, D^W) divides the sample space of the sliding window of residuals $r(k-W+1),...,r(k)$, or equivalently, the space of vectors of posterior probabilities, likelihood ratios, or log likelihood ratios of the sliding window of failure hypotheses into disjoint time-independent *sequential decision regions* $S_0, S_1, ..., S_N$. Here, $N = M$ if no failure time indication is involved in the terminal decision, while $N = MW$ if a failure time estimate is also required. Because the residuals are assumed to be Gaussian variables with variances that do not depend on the hypothesis, it is easy to check that an equivalent set of sufficient statistics is given by [20,23]

$$\Lambda(k) = [\Lambda'_0(k), ..., \Lambda'_{W-1}(k)]'$$

where for $0 \leq \sigma \leq W-1$

$$\Lambda_\sigma(k) = [\Lambda(k;1,\sigma),...,\Lambda(k;M,\sigma)]'$$

$$\Lambda(k;i,\sigma) = \sum_{s=0}^{\sigma} g_i'(s) V^{-1} r(k-\sigma+s)$$

Here σ indexes the possible failure onset times measured relative to the present time k (i.e. σ corresponds to a failure onset at time $k-\sigma$). The quantities $\Lambda(k;i,\sigma)$ differ only by an unimportant constant from the log-likelihood ratios for each hypothesis versus the no-fail hypothesis. The sliding window decision procedure operates as follows. At each time $k \geq W$, we form the decision statistics $\Lambda(k)$ from the window of residual samples. If $\Lambda(k) \in S_i$, for $i=1,...,N$, we stop sampling to declare δ_i; otherwise, $\Lambda(k) \in S_0$ and we proceed without making any immediate decision. The Bayes design problem is to determine a set of regions $S_0^*, S_1^*, ..., S_N^*$ that minimizes the corresponding sequential risk $U_s^W(\{S_i\})$ (the expression for which we will describe shortly). This represents a functional minimization problem that is generally very difficult to solve. A simplification of this problem is to constrain the decision regions to take on special shapes, $S_i(f)$, that are parameterized by a fixed dimensional vector f of design variables. A typical choice for these parametrically-specified regions might be in terms of the relative ordering of the sizes of the $L(k;i,\tau)$ and a set of threshold levels which correspond to the components of f (see (3) below). While such a constrained structure will lead to a suboptimal solution, the difference between the performance resulting from using the best constrained solution and that achieved by the optimal will be small if the constrained structure is chosen carefully. Furthermore, it is our contention that this performance difference will typically be mostly an artifact of the idealized problem formulation rather than a reality. That is, the unconstrained

problem seeks to find the best boundaries between decision regions, while the constrained problem fixes the boundary shapes (e.g. straight lines of polygonal boundaries). Given that the residual statistical model used to define the problem is subjected to error, the extra drop of performance resulting from being able to "fine tune" the boundary shapes will generally be dwarfed in the uncertainty arising from modeling errors.

In the remainder of this paper we focus our attention on a special set of parameterized sequential decision regions, because they are simple and they serve well to illustrate that the Bayes formulation can be exploited, in a systematic fashion, to obtain simple suboptimal rules that are capable of delivering good performance. These decision regions are:

$$S(j,t) = \{ \Lambda(k): \Lambda(k;j,t) > f(j,t),$$

$$\epsilon^{-1}(j,t)[\Lambda(k;j,t)-f(j,t)] > \epsilon^{-1}(i,s)[\Lambda(k;i,s)-f(i,s)], \quad (i,s) \neq (j,t) \} \tag{3a}$$

$$S(0,-) = \{ \Lambda(k): \Lambda(k;i,s) < f(i,s), \quad i=1,...,M, \; s=0,...,W-1 \} \tag{3b}$$

where $S(j,t)$ is the stop-to-declare-(j,k-t) region and $S(0,-)$ is the continue region. See Figure 1 for a pictorial representation of the structure of (3) in the case where there are only two failure hypothesized failure (j,k-t) and (i,k−s). Generally, the ϵ's may be regarded as design parameters, but here, $\epsilon(j,t)$ is simply taken to be the standard deviation of $\Lambda(k;j,t)$.

To evaluate $U_s^W(f)$, the Bayes risk due to the use of (3), we need to determine the set of probabilities, $Pr\{\Lambda(k) \in S(j,t), \Lambda(k-1) \in S(0,-),..., \Lambda(W) \in S(0,-) \mid i,\tau\}$, $k \geq W$, $j=1,...,M$, $t=0,...,W-1$, which, indeed, is the goal of many research efforts in so-called

level-crossing problems [24]. As it stands, each of the probabilities is an integral of a kMW-dimensional Gaussian density over the compound region $S(0,-) \times \cdots \times S(0,-) \times S(j,t)$, which, for large kMW, becomes extremely unwieldy and difficult to evaluate. A variety of approximations and bounds [25-28] have been developed for the evaluation of quantities such as this. We have not investigated the utility of any of these for our problem but rather have developed a systematic approach which is particularly appropriate for the dynamic FDI problem and which greatly simplifies the required calculations.

As a first step in this process, we reduce the dimension of the decision statistic $\Lambda(k)$ from MW to M. Specifically, we will base our decision process solely on the values of the log-likelihood ratios for each of the M failures modes assuming an onset time precisely at $\sigma = W-1$, i.e. the beginning of the window. Since we are not estimating failure time in this case, the terminal decision to be made is simply the identification of the failure modes. The rational behind this simplification has several aspects. First, in many applications, such as the aircraft sensor FDI problem [16] and the detection of freeway incidents [21], where the failure time need not be explicitly identifies, the failure time resolution power provided by the full window of decision statistics is not needed. Furthermore, even if failure onset time information is desired, resolution of this time within a block of length W may often be sufficient. If not, one can imagine a two-level decision-making structure in which one first determines the failure type (using the procedure to be described) and then estimate the onset time. Note that this overall system will have decidedly lower complicity than one based on simultaneous detection, identification and onset time estimation.

In assuming the utility of the approach just described one must make sure that the resulting decision algorithm does not have a significantly elevated probability of incorrectly identifying the failure type. That is, if a failure of type i occurs at a time before the end of the window and if a detection occurs, one would want the subsequent identification to also be i with high probability. Determining whether this is the case can be done completely in terms of the failure signatures [29]. We can expect good performance if cross-correlation among signatures for failures of the same type at different times are significantly higher than the cross-correlations of signatures corresponding to different failure types. We note that this is often the case in practice, and in fact an often-used goal for the residual generation process is that of producing signatures which are orthogonal or which at least lie in trivially overlapping subspaces [1,22].

A decision rule of the type just described consists of sequential decision regions that are similar to (3) but are only defined in terms of the M components $\Lambda(k;i,W-1)$, $i=1,\ldots,M$:

$$\Lambda_{W-1}(k) = [\Lambda(k;1,W-1), \Lambda(k;2,W-1),\ldots, \Lambda(k;M,W-1)]' \tag{4a}$$

$$S_j = \{ \Lambda_{W-1}(k): \Lambda(k;j,W-1) > f_j,$$

$$\epsilon^{-1}(j,W-1)[\Lambda(k;j,W-1)-f_j] > \epsilon^{-1}(i,W-1)[\Lambda(k;i,W-1)-f_i], \quad j\neq i \} \tag{4b}$$

$$S_0 = \{ \Lambda_{W-1}(k): \Lambda(k;j,W-1) \leq f_j, \quad j=1,\ldots M \} \tag{4c}$$

where S_j is the stop-to-declare-j region and S_0 is the continue region.

The risk for using (4) is

$$U_s^W(f) = L_F \sum_{i=1}^{M} \sum_{\tau=W+1}^{\infty} \mu(i,\tau) \sum_{k=W}^{\tau-1} \sum_{j=1}^{M} \Pr\{\Lambda_{W-1}(k) \in S_j, S_0(k-1) \mid 0,-\}$$

$$+ \sum_{i=1}^{M} \sum_{\tau=1}^{\infty} \mu(i,\tau) \sum_{k=\max[W,\tau]}^{\infty} \sum_{j=1}^{M} [c(i)(k-\tau)+L(i,j)] \Pr\{\Lambda_{W-1}(k) \in S_j, S_0(k-1) \mid i,\tau\}$$

where S_0 is the event defined below:

$$S_0(k) = \{ \Lambda_{W-1}(k) \in S_0, \ldots, \Lambda_{W-1}(W) \in S_0 \}$$

The first term in the expression for $U_s^W(f)$ represents the portion of the risk due to false alarms. The key expression here is $\Pr\{\Lambda_{W-1}(k) \in S_j, S_0(k-1) \mid 0,-\}$, which is the probability that no detections have been made before time k but that an identification for a type j failure is made a time k, given that no failure has occurred. The remaining portion of $U_s^W(f)$represents that part of the risk corresponding to detection delay and the possibility of incorrect identification. Here the key quantity is the probability $\Pr\{\Lambda_{W-1}(k) \in S_j, S_0(k-1) \mid i,\tau\}$, which is the probability that a detection is first made at time k and that the failure is identified as being of type j given that a type i failure occurred at time $\tau \leq k$. The calculation of these probabilities is specified by the following recursions:

$$\Pr\{\Lambda_{W-1}(k+1) \in S_j \mid S_0(k), i, \tau\}$$

$$= \left[\int_{S_0} p(\Lambda_{W-1}(k) \mid S_0(k-1), i, \tau)\, d\Lambda_{W-1}(k) \right]^{-1} x$$

$$\int_{S_0} p(\Lambda_{W-1}(k+1) \mid \Lambda_{W-1}(k), S_0(k-1), i, \tau)\, p(\Lambda_{W-1}(k) \mid S_0(k-1), i, \tau)\, d\Lambda_{W-1}(k),$$

$$k > W \qquad (5)$$

$$\Pr\{\Lambda_{W-1}(k) \in S_j, S_0(k-1) \mid i,\tau\}$$

$$= \Pr\{S_0(k-1) \mid 0,-\} \int_{S_j} p(\Lambda_{W-1}(k) \mid S_0(k-1),i,\tau) d\Lambda_{W-1}(k), \quad j=1,...,M \tag{6}$$

with

$$\Pr\{\Lambda_{W-1}(W) \in S_j \mid i,\tau\} = \int_{S_j} p(\Lambda_{W-1}(W) \mid i,\tau)\, d\Lambda_{W-1}(W) \tag{7}$$

Note that in essence what we are calculating in (5)-(7) are several different level crossing probabilities, and as we have just shown, it is these calculations that are the central elements to be determined in evaluating the performance of an hypothesized detection rule. For M small, numerical integration of (5)-(7) becomes manageable assuming that the required integrands are available.

Unfortunately, the transition density, $p(\Lambda_{W-1}(k+1) \mid \Lambda_{W-1}(k), S_0(k-1), i, \tau)$, required in (5) is difficult to calculate, because $\Lambda_{W-1}(k)$ is not a Markov process. In order to facilitate the computation of these probabilities, we use an approximation for this transition density obtained by developing an approximate Markovian model for the evolution of $\Lambda_{W-1}(k)$. A simple, but quite useful approximation is an an M-dimensional Gauss-Markov process $\ell(k)$ that is defined by

$$\ell(k+1) = A\,\ell(k) + \zeta(k+1) \tag{8a}$$

$$\text{cov}\{\zeta(k)\zeta'(t)\} = \Gamma\,\delta_{k,t} \tag{8b}$$

where A is an MxM constant matrix, and ζ is a white Gaussian sequence (with covariance equal to the (MxM) matrix Γ) uncorrelated with $\ell(k)$. The conditional mean of $\zeta(k)$ will be specified shortly. The reason for choosing this model is twofold.

First, just as $\Lambda_{W-1}(k)$, $I(k)$ is Gaussian. Second, $I(k)$ is Markov so that its transition density can be readily determined. In order to have the evolution of $I(k)$ match that of $\Lambda_{W-1}(k)$ as closely as possible, we choose the matrices A and Γ and the conditional mean $E_{i,\tau}\{\zeta(k)\}$ of $\zeta(k)$ under the hypothesis (i,τ) so that

$$E_{i,\tau}\{I(k)\} = E_{i,\tau}\{\Lambda_{W-1}(k)\} \tag{9a}$$

$$E_{0,-}\{I(k)I'(k)\} = E_{0,-}\{\Lambda_{W-1}(k)\Lambda'_{W-1}(k)\} \tag{9b}$$

$$E_{0,-}\{I(k)I'(k+1)\} = E_{0,-}\{\Lambda_{W-1}(k)\Lambda'_{W-1}(k+1)\} \tag{9c}$$

That is, we have matched the marginal density and the one-step cross-covariance of $I(k)$ to those of $\Lambda_{W-1}(k)$. A straightforward calculation shows that (8)-(10) uniquely specify

$$A = \Sigma'_1 \Sigma_0^{-1} \tag{10a}$$

$$\Gamma = \Sigma_0 - \Sigma'_1 \Sigma_0^{-1} \Sigma_1 \tag{10b}$$

$$E_{i,\tau}\{\zeta(k+1)\} = E_{i,\tau}\{\Lambda_{W-1}(k+1)\} - A\, E_{i,\tau}\{\Lambda_{W-1}(k)\} \tag{10c}$$

where

$$\Sigma_0 = E_{0,-}\{\Lambda_{W-1}(k)\Lambda'_{W-1}(k)\} = \sum_{t=0}^{W-1} G_t V^{-1} G'_t$$

$$\Sigma_1 = E_{0,-}\{\Lambda_{W-1}(k)\Lambda'_{W-1}(k+1)\} = \sum_{t=0}^{W-2} G_{t+1} V^{-1} G'_t$$

$$E_{i,\tau}\{\Lambda_{W-1}(k)\} = \begin{cases} 0, & \tau > k \\ \sum_{t=0}^{k-\tau} G_{t-k_0} V^{-1} g_i(t), & k_0 = k-W+1-\tau \leq 0 \\ \sum_{t=0}^{W-1} G_t V^{-1} g_i(t+k_0), & k_0 = k-W+1-\tau > 0 \end{cases}$$

$$G_t = [\, g_1(t), \ldots, g_M(t) \,]'$$

Clearly, Σ_0^{-1} exists if the failure signatures $[g'_i(0),\ldots,g'_i(W-1)]$, $i=1,\ldots,M$, are linearly independent. This condition is equivalent to the statement that there is sufficient information in a window of length W to distinguish among all of the M possible failure modes, assuming that if one of these failures has occurred, it did so in the beginning of the window. A sufficient condition for A to be stable, i.e. the magnitude of all its eigenvalues are less than unity, and Γ be invertible is that either G_0 or G_{W-1} is of rank M. (See the appendix for a discussion of the necessary and sufficient conditions for the invertibility of Γ and the stability of A.)

As an alternative to the model specification just given it is possible to choose other Markov approximations for $\Lambda_{W-1}(k)$. For example, one could match the n-step cross-covariance $(1<n<W)$ instead of matching the one-step cross-covariance as in (10). The suitability of a criterion for choosing the matrices A and Γ, such as (9) and (10), depends directly on the failure signatures under consideration and may be examined as an issue separate from the decision rule design problem. Also, a higher order Markov process may be used to approximate $\Lambda_{W-1}(k)$. However, the increase in the computational complexity may negate the benefits of the improved approximation. Finally, we emphasize that the statistics $l(k)$, as we have described it

here, is *not* an observable quantity. That is, it cannot be computed from the residuals. Rather, $l(k)$ is an artificial process introduced in order to obtain approximations for the calculation of the statistics of $\Lambda_{W-1}(k)$. Later in this section we will describe a suboptimal test statistics to replace $\Lambda_{W-1}(k)$ which is computable from the residuals and which is also Markov.

Using the model we have developed for $l(k)$ we can approximate the required probabilities by substituting $l(k)$ for $\Lambda_{W-1}(k)$ in the calculations. That is,

$$\Pr\{\Lambda_{W-1}(k)\in S_j, S_0(k-1)\,|\,i,\tau\}\approx\Pr\{l(k)\in S_j, S_0(k-1)\,|\,i,\tau\},\ j=0,1,\ldots,M,\ k>W$$

and

$$\Pr\{l(k)\in S_j, S_0(k-1)\,|\,i,\tau\} = \Pr\{S_0(k-1)\,|\,i,\tau\}\int_{S_j} p(l(k)\,|\,S_0(k-1)i,\tau)\,dl(k) \tag{11}$$

Assuming Γ^{-1} exists, we have

$$p(l(k+1)\,|\,S_0(k),i,\tau)$$

$$= \Big[\int_{S_0} p(l(k)\,|\,S_0(k-1),i,\tau)dl(k)\Big]^{-1} x$$

$$\int_{S_0}\big[p(\zeta(k+1)=(l(k+1)-A\,l(k))\,|\,i,\tau)p(l(k)\,|\,S_0(k-1),i,\tau)\big]\,dl(k) \tag{12}$$

where $p(\zeta(k)\,|\,i,\tau)$ is the Gaussian density of $\zeta(k)$ under the failure (i,τ). The key simplification that results from using the Markovian approximation is

$$p(l(k+1)\,|\,l(k),S_0(k-1),i,\tau) = p(l(k+1))\,|\,l(k),i,\tau)$$

$$= p(\zeta(k+1)=l(k+1)-A\,l(k)\,|\,i,\tau)$$

Because of this, the integrands in (12) are readily obtained (the first comes from the

previous step of the recursion) and thus the integrals in (12) can be calculated more easily.

In the event that Γ is not invertible, the density for $\zeta(k)$ is degenerate and (12) is more difficult to evaluate. As discussed in the appendix, the invertibility of Γ is related to the distinguishability of the M failure modes. Consequently, in any well-posed failure detection problem, W will be chosen so that the invertibility of Γ is assured.

Non-Window Sequential Decision Rules

Here we describe another simple decision rule that has the same decision region as the simplified sliding window rule (4), but the vector, z, of M-dimensional statistics is obtained differently as follows:

$$z(k+1) = \tilde{A}\, z(k) + \tilde{B}\, r(k+1) \tag{13}$$

where $\tilde{A}$ is a constant stable M x M matrix, and $\tilde{B}$ is a M x m constant matrix of rank M. Unlike the Markov model $l(k)$ that approximates $\Lambda_{W-1}(k)$, z(k) is a realizable Markov process driven by the residuals. The advantages of using z(k) as the decision statistics are: 1) less storage is required, because residual samples need not be stored as necessary in the sliding window scheme, and 2) since z(k) is Markov, the required probability integrals are of the forms (11) and (12) so that the same integration algorithm can be directly applied. Of course, z(k) is a suboptimal decision statistic. One could, if desired, use a higher-order model for z(k) so that it more nearly equals $\Lambda_{W-1}(k)$, but the added computational complexity may negate the advantages.

In order to form the statistics $z(k)$, we need to choose the matrices $\tilde{A}$ and $\tilde{B}$. When the failure signatures under consideration are constant biases and $M \leq m$, $\tilde{B}$ can simply be set to equal $G_0 V^{-1}$ (provided G_0 is of rank M), and $\tilde{A}$ can be chosen to be αI, where $0<\alpha<1$. Then, the term $\tilde{B}r$ in (13) provides the correlation of the residuals with the signatures as in (2), while the time constant α^{-1} characterizes the memory span of $z(k)$ just as W characterizes that of the sliding window statistics.

More generally, if we consider the case where failure signatures are not constant biases, $\text{rank}(G_0)<M$, or $m<M$, the choice of $\tilde{A}$ may still be handled in the same way as in the previous case, but the selection of $\tilde{B}$ is more involved. With some insights into the nature of the signatures, a reasonable choice of $\tilde{B}$ can often be made in order to have distinct components of $z(k)$ respond primarily to the corresponding failure. To illustrate how this may be accomplished, we will consider an example with two failure modes ($M=2$) and an m-dimensional residual vector. Let

$$g_1(k-\tau) = \beta_1$$

$$g_2(k-\tau) = \beta_2(k-\tau+1)$$

That is, g_1 is a constant bias, and g_2 is a ramp. If β_1 and β_2 are orthogonal a simple choice of $\tilde{B}$ is available:

$$\tilde{B} = \begin{bmatrix} \beta'_1 \\ \beta'_2 \end{bmatrix}$$

This choice may often be acceptable even when $\beta'_1\beta_2 \neq 0$. It is clearly not of any use when β_1 and β_2 are multiples of the same vector β, or when they are scalars (corresponding to $m=1$), as the rank of $\tilde{B}$ is less than 2. In these cases we can

consider processing groups of residuals. For example, suppose we batch process every two residual samples together, i.e. we use the augmented residual sequence $\tilde{r}(k) = [r'(2\tilde{k}-1), r'(2\tilde{k})]'$, $\tilde{k}=1,2,....$ In this case we can set $\tilde{B}$ to be

$$\tilde{B} = \begin{bmatrix} \beta' & \beta' \\ \beta' & 2\beta' \end{bmatrix}$$

Thus, this $\tilde{B}$ is of dimension M x 2m and has rank M (=2). The first and second rows of $\tilde{B}$ captures the constant bias and ramp nature g_1 and g_2 , respectively. The use of the modified residual $\tilde{r}(k)$ in this case causes no adverse effect, since it only lengthens slightly the interval between times when terminal decisions can be made. Clearly one can consider further augmentation and batch processing of the residuals, and in general the logical choice of $\tilde{B}$ is one in which each row of $\tilde{B}$ contain in sequence the initial values of the corresponding failure signature. In this case the mean values of z(k) will exactly equal that of $\Lambda_{W-1}(k)$ for a number of time steps following a failure equal to the level of augmentation used. The utility of this approach clearly depends on the temporal structure of the failure signatures. For problems where the signatures vary drastically as a function of the elapsed time and the distinguishability among failures depends essentially on these variations, the effectively of using z(k) diminishes. In such cases the sliding window decision rule should provide better performance, although it should be noted that in this case one would typically have to use a comparatively long window in order to obtain an adequate degree of distinguishability.

3.2 Risk Evaluation

An algorithm based on 1-dimensional Gaussian quadrature formulas [30,31] has been developed to compute the probability integrals of (11) and (12) for the case $M=2$. (It can be extended to higher dimension with an increase in computation.) The details of this quadrature algorithm is described in [22]. Its accuracy has been assessed via comparison with Monte Carlo simulations (see the numerical example in Section 4). With this algorithm we can evaluate the performance probabilities and risks associated with the suboptimal decision rules described above.

In the absence of a failure, the conditional density for $I(k)$ (12) has been observed in numerous examples to essentially reach a steady-state at some finite time $T > W^*$. Assuming this is the case, we have for $k \geq \tau \geq T$,

$$\Pr\{I(k) \in S_j \mid S_0(k-1), 0, -\} = \bar{b}_j \tag{14}$$

$$\Pr\{I(k) \in S_j, I(k-1) \in S_0, \ldots, I(\tau) \in S_0 \mid S_0(\tau-1), i, \tau\} = b_j(k-\tau \mid i) \tag{15}$$

That is, once steady-state is reached, only the elapsed time since failure is important. Generally, failures occur infrequently, and decision rules with low false alarm probability are employed. Thus, it is reasonable to assume 1) $\rho \ll 1$, i.e. $(1-\rho)^{-1} \approx 1$, and 2) $\Pr\{S_0(T) \mid 0, -\} \approx 1$. The sequential risk associated with (4) for $M=2$ can be approximated by

$$U_s^W(f) = P_F L_F + (1-P_F) \sum_{i=1}^{2} a(i) \sum_{j=1}^{2} \sum_{t=0}^{\infty} [c(i) + L(i,j)]\, b_j(t \mid i) \tag{16}$$

where

---

* Unfortunately, we have not been able to prove such convergence behavior using elementary techniques. More advance function-theoretic methods may be necessary.

$$P_F = \frac{(1-\rho)(1-\bar{b}_0)}{1-\bar{b}_0(1-\rho)}$$

P_F is the unconditional false alarm probability, i.e. the probability of one false alarm over all time.

Next, we seek to replace the infinite sum over t in (16) by the finite sum up to t=Q plus a term approximating the remainder of the infinite sum. Suppose we have been sampling for Q steps since a failure occurred. Define

$$P_t(j|i) = \Pr\{I(t) \in S_j | S_0(t-1), i, 0\}, \quad j=0,1,2$$

If we stop computing the probabilities after Q, we may approximate

$$P_t(j|i) \approx P_Q(j|i), \quad j=0,1,2, \quad t>Q \tag{17}$$

That is we assume that after a detection delay of Q steps the conditional probability of detection at any time given *no* detection at any previous time reaches a constant steady-state value. This is the same as assuming that beyond Q steps of delay, the additional detection delay is exponentially distributed. This assumption is reasonable for constant failure signatures or signatures that reach steady-state. While the assumption may not be valid for signatures which continue to vary, the effect of this approximation is generally quite small, since for all i and j one typically can choose Q so that

$$b_j(t|i) \approx 0, \quad t>Q$$

That is, for each failure mode the probability of a detection delay greater than Q steps is negligible.

Substituting (17) in (16), we obtain

$$U_s^W(f) = P_F L_F + (1-P_F)\sum_{i=1}^{2} a(i)\left[c(i)\bar{t}_i + \sum_{j=1}^{2} L(i,j)P(i,j)\right] \tag{18}$$

where

$$\bar{t}_i = \sum_{j=1}^{2}\sum_{t=0}^{Q} t b_j(t|i) + Q b_0(Q|i) + \frac{1}{1-P_Q(0|i)} \tag{19}$$

$$P(i,j) = \sum_{t=0}^{Q} b_j(t|i) + b_0(Q|i)\frac{P_Q(j|i)}{1-P_Q(0|i)} \tag{20}$$

Here, $\bar{t}_i$ is the conditional expected delay to decision, given that a type i failure has occurred, and $P(j|i)$ is the conditional probability of declaring a type j failure, given that failure i has occurred. From the assumption that $Pr\{S_0(T)|0,-\}\approx 1$ and the steady-state condition (14), it can be shown that the mean time between false alarms is simply $(1-\bar{b}_0)^{-1}$. Now all the probabilities in (18)-(20) can be computed by using our quadrature algorithm. Note that the risk expression (18) consists only of finite sums and it can be evaluated with a reasonable amount of computational effort. With such an approximation of the sequential risk, we are able to consider the problem of determining the decision regions (i.e. the thresholds f_j's) that minimizes the risk.

It should be noted that we could consider choosing a set of thresholds that minimizes a weighted combination of certain detection probabilities ($P(i,j)$), the expected delay ($\bar{t}_i$), and the mean time between false alarms $(1-\bar{b}_0)^{-1}$. Although such an objective function will not result in a Bayesian design in general, it is a valid design criterion that may be useful for some applications.

3.3 Risk Minimization

The risk minimization has two features that deserves special attention. First, the sequential risk is not a simple function of the threshold f, and its derivatives with respect to f is not readily available. Second, calculating the risk is a computationally intensive task. Therefore, the minimum-seeking procedure to be used must require few function evaluations, and it must not require derivatives. For these reasons we chose to use the Sequence-of-Quadratic-Programs (SQP) algorithm studied by Winfield [32] to solve this problem, because it does not need any derivative information and it appears to require fewer function evaluations than other well-known algorithms [32]. Furthermore, the SQP is simple, and it has quadratic convergence. Very briefly, the algorithm consists of the following. At each step of the iteration, a quadratic surface is fitted to the risk function locally using the preceding guesses at the optimal value of f and the corresponding risk function evaluations. The resulting quadratic model is minimized over a constrained region (hence the name SQP). The risk function is evaluated at this minimum and is used in the surface fitting of the next iteration. The details of the application of SQP to risk minimization is reported in [22].

4. NUMERICAL EXAMPLE

Now we discuss an application of the suboptimal rule design methodology developed in this paper. We consider the detection of two possible failure modes (without identifying the failure time). The residual is a 2-dimensional vector, and the vector failure signatures, $g_i(t)$, $i=1,2$, as functions of the elapse time t are shown in Table 1. The signature of the first failure is simply a constant vector. The first component of $g_2(t)$ is a constant, while the second component is a ramp. We have chosen to examine these those types of signatures because they are simple and describe a large variety of failure signatures that are commonly seen in practice. For simplicity, we have chosen V, the covariance of r, to be the identity matrix.

Both a simplified sliding window rule (that uses Λ_{W-1}) and a rule using the Markov statistic z were examined. The parameters associated with Λ_{W-1}, l, and z are shown in Table 2, and the cost functions and the prior probability are shown in Table 3. To facilitate discussion, we introduce the following terminology. We refer to a Monte Carlo simulation of the sliding window rule by SW, a simulation of the rule using the Markov statistic z as Markov Implementation (MI), and a simulation of the non-implementable decision process using the approximation l as Markov Approximation (MA). (All simulations are based on 10,000 trajectories.) The notation Q20 refers to the results of applying the quadrature algorithm to calculate the various performance indices of the sliding window rule while using l to approximate Λ_{W-1} (12).

The results of SW, MA, and Q20 for the thresholds [8.85, 12.05] are shown in Figures 2-6 (see (15) for the definition of notation). The quadrature results Q20 are

very close to those obtained by Monte Carlo simulations for MA, indicating the excellent accuracy of the quadrature algorithm. In comparing SW with MA, it is evident that the Markov approximation slightly under-estimates the false alarm rate of the sliding window rule (SW). However, the response of the Markov approximation to failure is very close to that of the sliding window rule. In the present example, Λ_{W-1} is a 7-th order process, while its approximation l is only of first order. In view of this fact we can conclude that l provides a very reasonable and useful approximation of Λ_{W-1}.

The successive choices of thresholds by SQP for the sliding window rule are plotted in Figure 7. Note that we have not carried the SQP algorithm so far that the successive choices of thresholds are, say, within .001 of each other. This is because near the optimum the expected risk is relatively insensitive to small changes in f. This implies that fine scale optimization is not generally worthwhile. This conclusion is supported by the fact that the residual signature models used in designing failure detection systems are typically idealizations, and thus minor improvements in Bayes risk is generally an artifact of the mathematical formulation. Furthermore, it should be remembered that the use of the Bayes formulation is simply for the purpose of providing a mechanism for determining high-performance decision rules, and thus the precise optimization of the Bayes risk is not the central issue. In fact, the cost parameters L, c, μ, and W should be used as design parameters. In the event that the optimal thresholds resulting from a particular choice of Bayes risk do not provide the desired detection performance, the design parameters may be adjusted and the SQP may be repeated to get a new design. A practical alternative method is to make use of

the list of performance indices (e.g. P(i,j)) that are generated in the risk calculation, and choose a pair of thresholds that yields the desired performance tradeoff.

The performance of the decision rule using Λ_{W-1} and z as determined by SQP are shown in Figures 8-12. (The thresholds for Λ_{W-1} are [8.85, 12.05] and those for z are [6.29, 11.69].) We note that MI has a higher false alarm rate than SW. The speeds of detection for the two rules are similar. While MI has a slightly higher type 1 correct detection probability ($\sum_{s=0}^{t} b_1(s|1)$) than SW, SW has a consistently higher type 2 correct detection probability ($\sum_{s=0}^{t} b_2(s|2)$ than MI. By raising the thresholds of the rule using z appropriately, we can decrease the false alarm rate of MI down to that of SW with an increase in detection delay and slightly improved correct detection probability for the type 2 failure. Thus the sliding window rule is slightly superior to the rule using z in the sense that when both are designed to yield a comparable false alarm rate, the latter will have longer detection delays and a slightly lower correct detection probability for a type 2 failure. In view of the fact that a decision rule using z is much simpler to implement, it is worthy of being considered as an alternative to the sliding window rule.

In summary, this example illustrates the utility of our approach. The quadrature algorithm has been shown to be accurate and useful, and the Markov approximation of Λ_{W-1} by l is a valid one. The simplicity and usefulnes of the SQP algorithm have also been demonstrated. Finally, the Markov decision statistic z has been shown to be a worthy alternative to the sliding window statistic Λ_{W-1}.

5. CONCLUSION

A computationally feasible methodology based on the Bayesian approach has been developed for designing suboptimal sequential decision rules for FDI. This methodology was applied to a numerical example, and the results indicate that it is a potentially useful design approach.

APPENDIX

Theorem

Consider the Gauss-Markov process $\Lambda(k)$ specified by (8)-(10) in Section 3.1. Given $E\{\Lambda_{W-1}(k)\Lambda'_{W-1}(k)\} = \Sigma_0 > 0$, A has at least one unity eigenvalue and Γ is semi-positive definite if and only if there exist M-vectors $\alpha \neq 0$ and $\beta \neq 0$ such that

$$\alpha' G_i = \beta' G_{i+1}, \quad i=0,\ldots,W-2 \tag{A1}$$

$$\alpha' G_{W-1} = 0 \tag{A2}$$

and

$$\beta' G_0 = 0 \tag{A3}$$

Proof

Let

$$\tilde{\Sigma} = E\left\{\begin{bmatrix} \Lambda_{W-1}(k) \\ \Lambda_{W-1}(k+1) \end{bmatrix} [\Lambda'_{W-1}(k), \Lambda'_{W-1}(k+1)]\right\} = \begin{bmatrix} \Sigma_0 & \Sigma_1 \\ \Sigma'_1 & \Sigma_0 \end{bmatrix}$$

Using the transformation T

$$T = \begin{bmatrix} I & 0 \\ -\Sigma'_1\Sigma_0^{-1} & I \end{bmatrix}$$

we obtain

$$T\tilde{\Sigma}T' = \begin{bmatrix} \Sigma_0 & 0 \\ 0 & \Gamma \end{bmatrix}$$

Since T is full rank and $\Sigma_0 > 0$, $\tilde{\Sigma}$ and Γ are semi-positive definite if and only if there are non-zero M-vectors α and β such that

$$\alpha' \Lambda_{W-1}(k+1) = \beta' \Lambda_{W-1}(k) \tag{A4}$$

Recall

$$\Lambda_{W-1}(k) = \sum_{s=0}^{W-1} G_s V^{-1} r(k-W+1+s)$$

Therefore, (A4) is equivalent to the conditions (A1)-(A3). From (8), we obtain

$$\Sigma_0 = A \Sigma_0 A' + \Gamma$$

It follows that A only has eigenvalues of magnitudes less than or equal to unity, and it has at least one unity eigenvalue if and only if Γ is semi-positive definite.

Q. E. D.

Suppose all signatures vanish for elapse times greater than W-1, i.e. $g_i(t)=0$, for $t > W-1$, and $i=1,...M$. Then, (A1)-(A3) are equivalent to the condition that it is not possible to distinguish between a failure occurring at a certain time and failures occurring one time step earlier or later. Moreover, (A1)-(A3) indicate that only a special class of failure signatures would satisfy this indistinguishability condition for all value of W. Generally, it is possible to choose a sufficiently large W so that this situation is avoided.

$$g_1(t) = \begin{bmatrix} 1 \\ .5 \end{bmatrix}, \qquad g_2(t) = \begin{bmatrix} .5 \\ .25 + .25t \end{bmatrix}, \qquad V = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Table 1. Failure Signatures.

$$W = 8$$

$$\Sigma_0 = \begin{bmatrix} 10 & 8.5 \\ 8.5 & 14.75 \end{bmatrix}, \qquad A = \begin{bmatrix} .826 & .058 \\ .116 & .837 \end{bmatrix}, \qquad \Gamma = \begin{bmatrix} 2.32 & 2.01 \\ 2.01 & 4.58 \end{bmatrix}$$

$$\tilde{A} = \begin{bmatrix} .875 & 0 \\ 0 & .875 \end{bmatrix}, \qquad \tilde{B} = \begin{bmatrix} 1 & .5 \\ .5 & 2 \end{bmatrix}$$

Table 2. Parameters for Λ_{W-1}, l, and z.

$$c_1 = c_2 = 1$$

$$L(1,2) = L(2,1) = 10, \quad L(1,1) = L(2,2) = 0, \quad L_F = 9$$

$$T = 8, \qquad Q = 8, \qquad \rho = .0002$$

$$\mu(i,\tau) = .5\rho(1-\rho)^{\tau-1}, \quad i = 1,2$$

Table 3. Cost Functions and Prior Probability.

REFERENCES

[1] E. Y. Chow and A. S. Willsky, "Analytical redundancy and the design of robust failure detection systems," submitted to *IEEE Trans. Automat. Contr.*

[2] A. Wald, *Sequential Analysis*, Wiley, New York, 1947.

[3] D. Blackwell and M. A. Girshick, *Theory of Games and Statistical Decisions*, Wiley, New York, 1951.

[4] H. Chernoff, *Sequential Analysis and Optimal Design*, SIAM Regional Conference Series in Applied Mathematics, vol. 8, Philadelphia, 1972.

[5] E. S. Page, "Continuous inspection schemes," *Biometrika*, vol. 41, pp. 100-114, 1954.

[6] R. M. Gagliardi and I. S. Reed, "On the sequential detection of emerging targets," *IEEE Trans. Inform theory*, vol. IT-11, pp. 260-263, April 1965.

[7] T. T. Chien and M. B. Adams, "A sequential failure detection technique and its application," *IEEE Trans. Automat. Contr.*, vol. AC-21, pp. 750-757, Oct. 1976..

[8] M. Basseville, "Edge detection using sequential methods for change in level - Part II: Sequential detection of change in mean," *IEEE Trans. Acoust. Speech, Signal Processing*, vol. ASSP-29, pp. 32-50, Feb. 1981.

[9] A. N. Shiryaev, "On optimum methods in quickest detection problems," *Theory of Probability and Applications*, vol. 8, pp. 22-46, 1963.

[10] A. N. Shiryaev, "On the detection of disorder in a manufacturing process I," *Theory of Probability and Applications*, vol. 8, pp.247-265, 1963.

[11] A. N. Shiryaev, "On the detection of disorder in a manufacturing process II," *Theory of Probability and Applications*, vol. 8, pp.402-413, 1963.

[12] R. A. Roberts and C. T. Mullis, "A Bayes sequential test of M hypotheses," *IEEE Trans. Inform. Theory*, vol. IT-16, pp. 91-94, Jan. 1970.

[13] A. H. El-Sawy and V. D. Vandelinde, "Robust sequential detection of signals in noise," *IEEE Trans. Inform. Theory*, vol. IT-25, pp. 346-353, May 1979.

[14] S. Fleisher and E. Shwedyk, "A sequential multiple hypotheses test for the unknown parameters of a Gaussian distribution," *IEEE Trans. Inform. theory*, vol. IT-26, pp. 255-259, March, 1980.

[15] T. S. Ferguson, *Mathematical Statistics, A Decision Theoretic Approach*, Academic Press, New York, 1967.

[16] J. C. Deckert, M. N. Desai, J. J. Deyst, and A. S. Willsky, "F-8 DFBW sensor failure identification using analytic redundancy," *IEEE Trans. Automat. Contr.*, vol. AC-22, pp. 795-803, Oct. 1977.

[17] D. E. Gustafson, A. S. Willsky, J. Y. Wang, M. C. Lancaster, and J. H. Triebwasser, "ECG/VCG rhythm diagnosis using statistical signal analysis - II. Identification of transient rhythms," *IEEE Trans. Biomed. Eng.*, vol. BME-25,

pp. 353-361, July, 1978.

[18] R. J. McAulay and E. Denlinger, "A decision-directed adaptive tracker," " *IEEE Trans. Aeros. and Elec. Sys.*, vol. AES-9, pp. 229-236, March, 1973.

[19] J. Korn, S. W. Gully, and A. S. Willsky, "Application of the generalized likelihood ratio algorithm to maneuver detection and estimation," *Proc. American Control Conference*, 1982.

[20] A. S. Willsky and H. L. Jones, "A generalized likelihood ratio approach to the detection and estimation of jumps in linear systems," *IEEE Trans. Automat. Contr.*, vol. AC-21, pp. 108-112, Feb. 1976.

[21] A. S. Willsky, E. Y. Chow, S. B. Gershwin, C. S. Greene, P. K. Houpt, and A. L. Kurkjian, "Dynamic model-based techniques for the detection of incidents on freeways," *IEEE Trans. Automat. Contr.*, vol. AC-25, pp. 347-360, June 1980.

[22] E. Y. Chow, *A Failure Detection System Design Methodology*, SC.D. Thesis, Laboratory for Information and Decision Systems, M.I.T., Cambridge, MA, 1980.

[23] H. L. Van Trees, *Detection, Estimation, and Modulation Theory, Vol. I*, Wiley, New York, 1968.

[24] I. F. Blake and W. C. Lindsey, "Level-crossing problems for random processes," *IEEE Trans. Inform. Theory*, vol. IT-19, pp.295-315, May 1973.

[25] R. G. Gallagher and C. W. Helstrom, "A bound on the probability that a Gaussian process exceeds a given function," *IEEE Trans. Inform Theory*, vol. IT-15, pp. 163-166, Jan. 1969.

[26] D. H. Bhati, "Approximations to the distribution of sample size for sequential tests, I. Tests of simple hypotheses," *Biometrika*, vol. 46, pp. 130-138, 1973.

[27] B. K. Ghosh, "Moments of the distribution of sample size in a SPRT," *American Statistical Assoc. Journal*, vol. 64, pp. 1560-1574, 1969.

[28] B. K. Walker, *A Semi-Markov Approach to Quantifying Fault-Tolerant System Performance*, Sc.D. Thesis, Dept. of Aero. and Astro., M.I.T., Cambridge, MA, 1980.

[29] E. Y. Chow, *Analytical Studies of the Generalized Likelihood Ratio Technique for Failure Detection*, S.M. Thesis, Dept. of Elec. Eng. and Comp. Sci., M.I.T., Cambridge, MA, 1976.

[30] P. J. Davis and P. Rabinowitz, *Numerical Integration*, Blaisedell, Waltham, MA, 1967

[31] A. H. Stroud and D. Secrest, *Gaussian Quadrature Formulas*, Prentice Hall, Englewood Cliffs, N. J., 1966.

[32] D. H. Winfield, *Function Minimization Without Derivatives by a Sequence of Quadratic Programming Problems*, Technical Report No. 537, Division of Eng. and Applied Physics, Harvard Univ., Cambridge, MA, Aug. 1967.

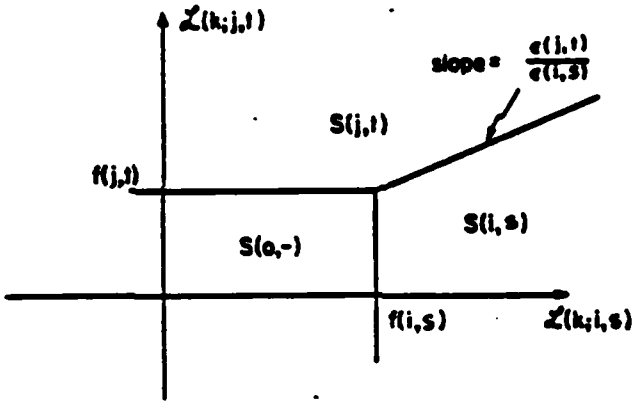


Figure 1. Sequential Decision Regions in 2 Dimensions

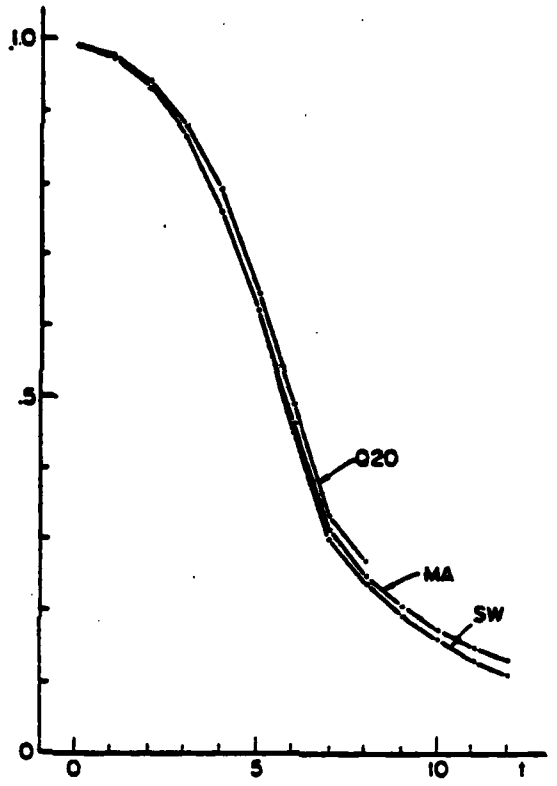


Figure 3. $b_0(t \mid 1)$: SW, MA, and Q20

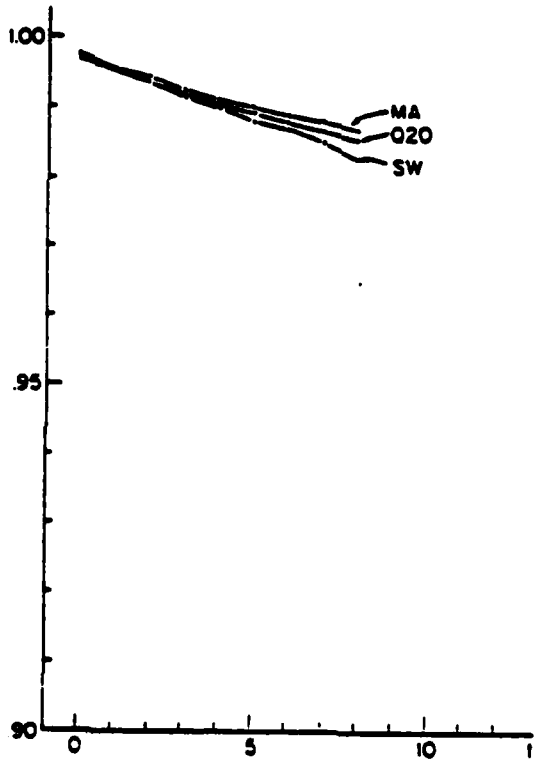


Figure 2. $b_0(t \mid 0)$: SW, MA, and Q20

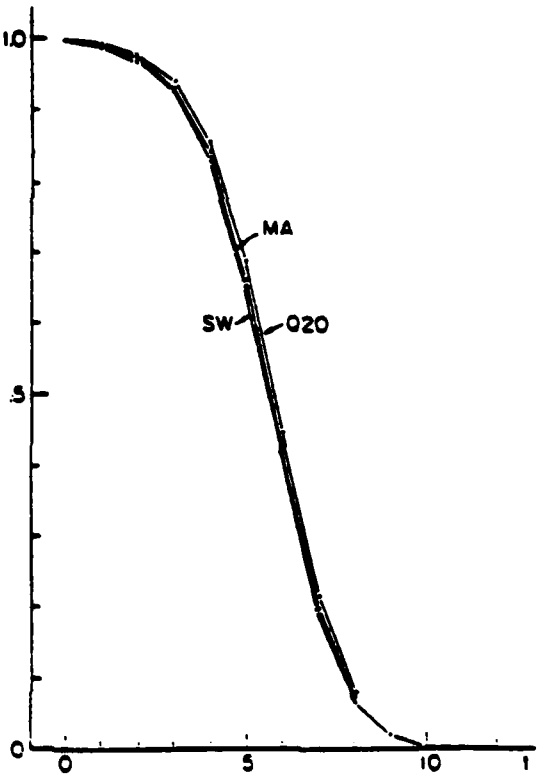


Figure 4. $b_0(t \mid 2)$: SW, MA, and Q20

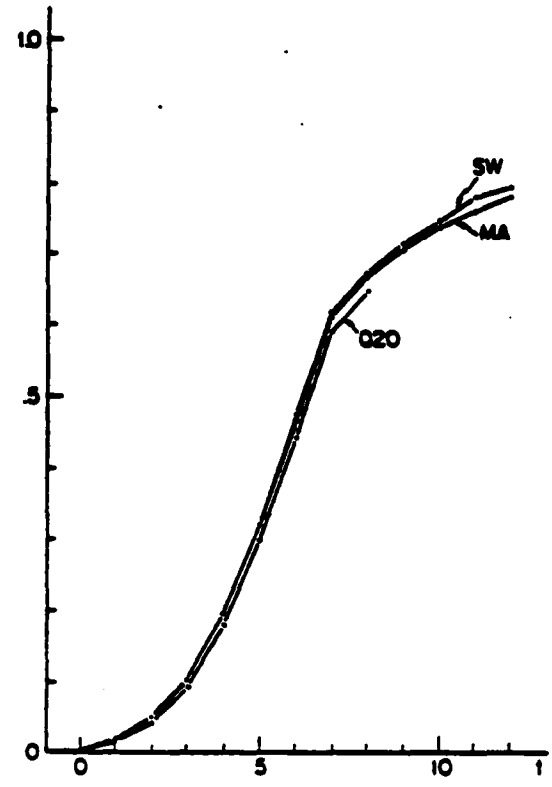


Figure 5. $\sum_{s=0}^{t} b_1(s \mid 1)$: SW, MA, and Q20

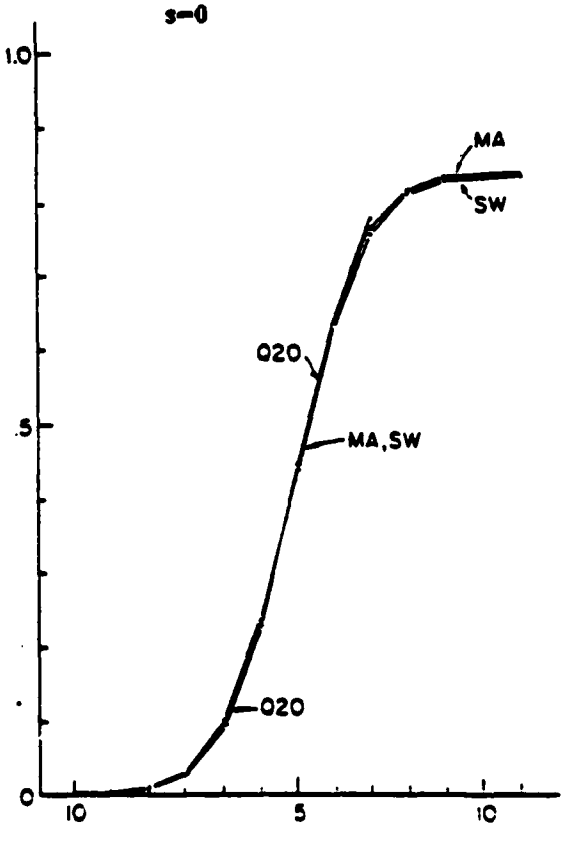


Figure 6. $\sum_{s=0}^{t} b_2(s \mid 2)$: SW, MA, and Q20

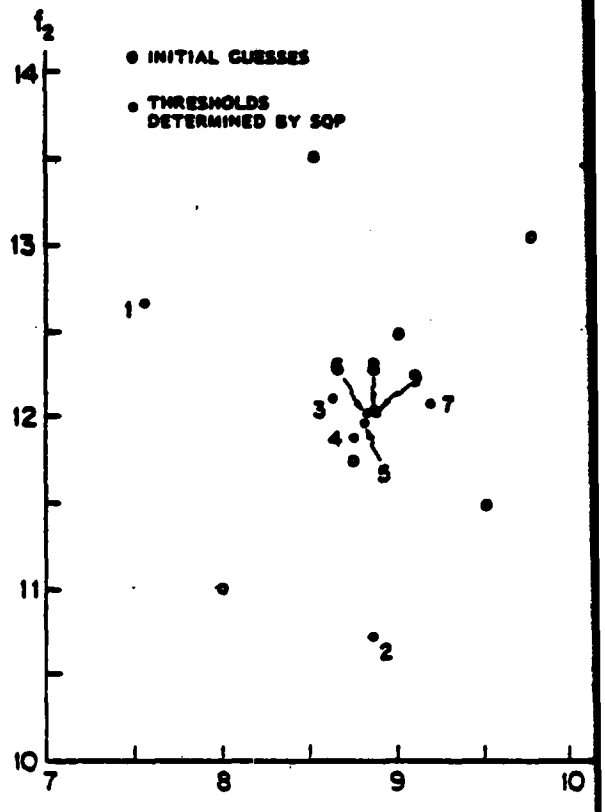


Figure 7. Thresholds Chosen by SQP

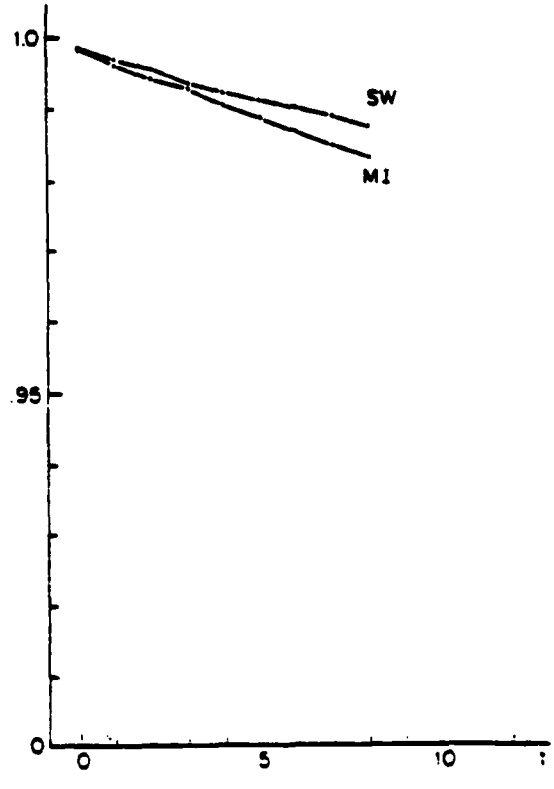


Figure 8. $b_0(t \mid 0)$: SW and MI

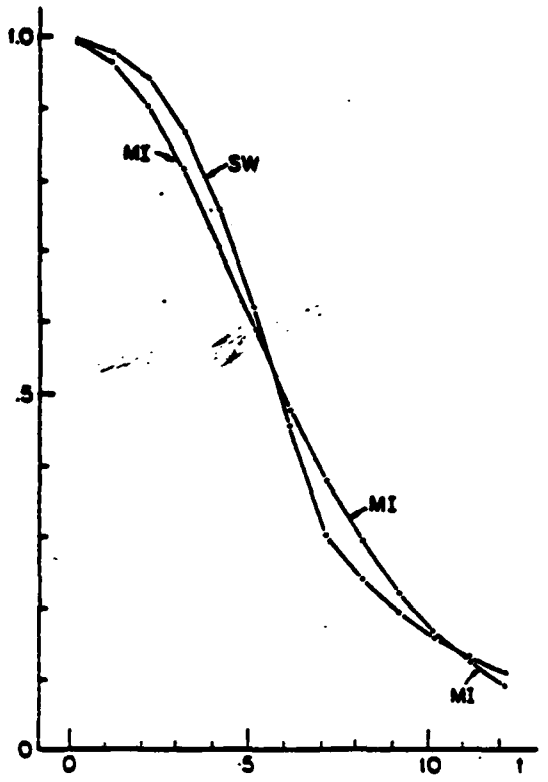


Figure 9. $b_0(t \mid 1)$: SW and MI

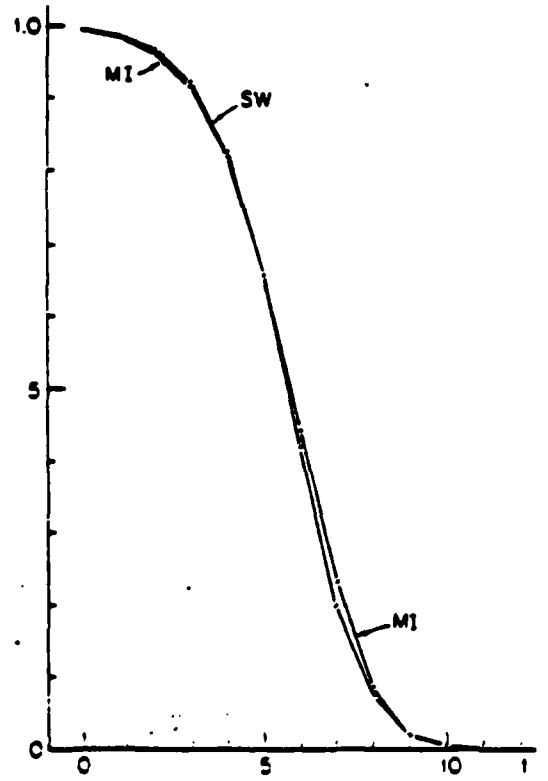


Figure 10. $b_0(t \mid 2)$: SW and MI

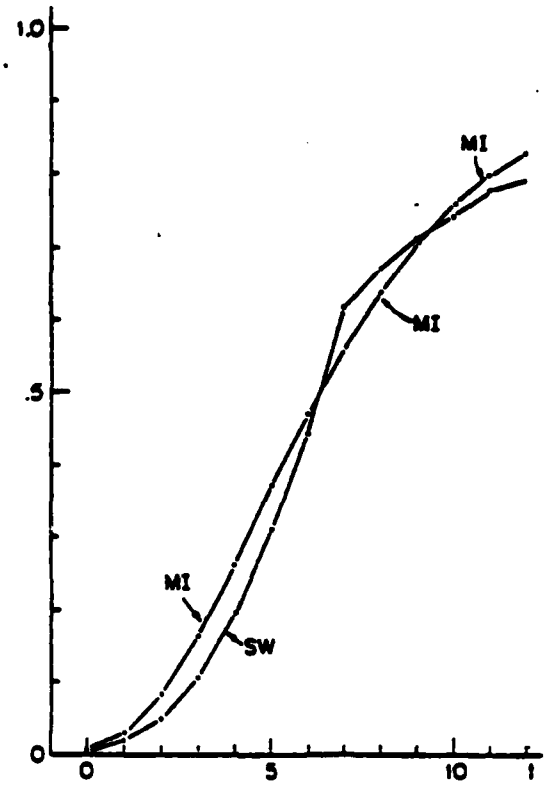


Figure 11. $\sum_{s=0}^{t} b_1(s \mid 1)$: SW and MI

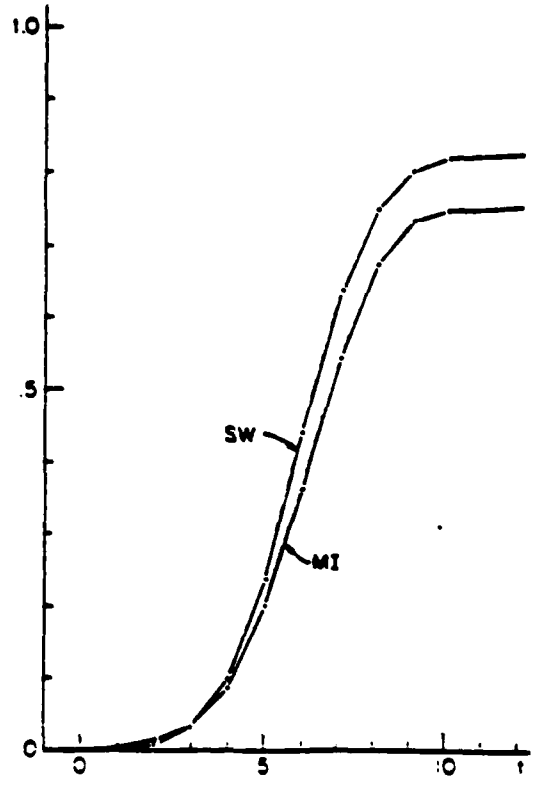


Figure 12. $\sum_{s=0}^{t} b_2(s \mid 2)$: SW and MI

END

END

FILMED

5-83

DTIC